

## Data Elements

The data in the Data Warehouse comes from multiple sources including OSIRIS, SASI, FIPPS, various test vendors, and the State. Some assumptions were made during the design processes that were found to be invalid during the load process due to incorrect data in the source systems. Dealing with incorrect data has been difficult but the majority of the information has been accommodated or corrected. If you encounter data that you believe to be invalid, please notify the Data Warehouse team so that the data can be investigated and corrected.

We will produce a more comprehensive data dictionary in the future. For now, this is a summary of the data that is loaded into the Data Warehouse and potential issues that may exist for a very small percentage of the data.

### ***School / School Year Information***

	'97-'98	'98-'99	'99-'00	'00-'01	'01-'02*
<b>School Year Data – Manual</b>	X	X	X	X	X
<b>School Data – Manual</b>					X

**Issues:** School and school year information has been captured manually and populated in the Data Warehouse. This data will be maintained as new data is loaded. The current design does not allow for historical capture of the information so all information (such as cluster information) is stored as it currently exists. We intend to address this issue in a future phase.

### ***Staff Information***

	'97-'98	'98-'99	'99-'00	'00-'01	'01-'02*
<b>Staff Data – FIPPS</b>	X	X	X	X	X

\*'01-'02 data is as of March 8, 2002

**Issues:** The current design does not allow for historical capture of the information so all information (such as job class information) is stored as it currently exists. We intend to address this issue in a future phase.

### ***Student / Enrollment Information***

	'97-'98	'98-'99	'99-'00	'00-'01	'01-'02*
<b>Elementary Data – SASI</b>	X	X	X	X	X
<b>Middle School Data – OSIRIS</b>	X	X			
<b>Middle School Data – SASI</b>			X	X	X
<b>High School Data – OSIRIS</b>	X	X			
<b>High School Data - SASI</b>			X	X	X

\*'01-'02 data is as of March 8, 2002

**Issues:** Source data contained duplicate Social Security Numbers assigned to different students (sometimes siblings and sometimes totally different students). We also encountered situations where a student had multiple Perm Numbers assigned, and a few

## Data Elements

cases where the same Perm Number was assigned to two students. In some cases, the transformation processes were confused by these situations and data for these students may be suspect. These students only represent approximately 0.2% of the total student population. We are continuing to investigate and make corrections to these records.

### ***CRCT Test Data***

	<i>Spr '98</i>	<i>Spr '99</i>	<i>Spr '00</i>	<i>Spr '01</i>	<i>Spr '02</i>
<b>CRCT Data – Riverside Publishing</b>			X		
<b>CRCT Data – Advanced Measurement</b>				X	

**Issues:** There were different vendors for each year so some of the ancillary information is not available for the Spring 2000 test. Much of the data was miscoded and matching to students was sometimes difficult. As a result, a few of the test records have not been loaded since we have not been able to tie the tests to a student.

### ***HSGT Test Data***

	Fall '97	Win '97	Spr '98	Sum '98	Fall '98	Win '98	Spr '99	Sum '99	Fall '99	Win '99	Spr '00	Sum '00	Fall '00	Win '00	Spr '00
<b>General</b>	X	X	X	X	X	X	X	X		X	X	X	X	X	X
<b>Writing</b>	X	n/a	X	X	X	n/a	X	X		n/a	X	X	X	n/a	X

**Issues:** Note that the Fall of 1999 data is missing. It has been ordered and will be loaded after it is received. The Writing test is not administered in the winter session. Like the CRCT data, some of the data was miscoded and was very difficult to match to students. Some of the test records have not been loaded since we have not been able to tie the tests to a student. We have also encountered situations where students have taken the test multiple times in the same test period. This data has been accommodated by showing the test dates as consecutive days.

### ***Attendance Information***

	'97-'98	'98-'99	'99-'00	'00-'01	'01-'02*
<b>Elementary Data – SASI</b>	X	X	X	X	X
<b>Middle School Data – OSIRIS</b>	X	X			
<b>Middle School Data – SASI</b>			X	X	X
<b>High School Data – OSIRIS</b>	X	X			
<b>High School Data - SASI</b>			X	X	X

\*'01-'02 data is as of March 8, 2002

**Issues:** Attendance information is captured on a total school/school year basis. If a student is enrolled in a school multiple times in a school year (i.e. is enrolled, withdraws, and comes back later in the same year), the attendance information represents the total number of days present or absent over all of the attendance periods. We cannot tell how

## Data Elements

many days they were present or absent for each enrollment period. We intend to address this issue in a future phase.

### *Course / Mark Information*

	'97-'98	'98-'99	'99-'00	'00-'01	'01-'02*
<b>Elementary Data – SASI</b>	X	X	X	X	X
<b>Middle School Data – OSIRIS</b>	X	X			
<b>Middle School Data – SASI</b>			X	X	X
<b>High School Data – OSIRIS</b>	X	X			
<b>High School Data - SASI</b>			X	X	X

\*'01-'02 data is as of March 8, 2002

**Issues:** While individual course offering information (by course, teacher, and period) is stored in the Data Warehouse, the students cannot be tied to a specific course offering due to limitations in the current sourcing process. All grades are tied to the correct teacher, semester, and student, but not the correct period when a teacher teaches the same course multiple times in a day. We intend to address this issue in future phase.

### *SAT Test Data*

	<i>1997</i>	<i>1998</i>	<i>1999</i>	<i>2000</i>	<i>2001</i>
<b>SAT Data</b>	X	X	X	X	

SAT test data was provided by the testing vendors for 1997 – 2000 school years.

Much of the data was miscoded and matching to students was sometimes difficult. As a result, a few of the test records have not been loaded since we have not been able to tie the tests to a student.